

Comment on essd-2021-182

Anonymous Referee #2

Referee comment on "Historical reconstruction of background air pollution over France for 2000–2015" by Elsa Real et al., Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2021-182-RC2, 2021

The paper by Real et al. proposes a dataset of background air pollution concentrations and air quality indicators over France for the period 2000–2015. The concentrations and indicators are mainly given on an annual basis either gridded at about 4km resolution or aggregated on French administrative territories. The provided concentrations are calculated using kriging approaches merging surface measurements from air quality networks and model simulation. The evaluation of the dataset is done using a crossvalidation method and shows good performances of the dataset to assess air pollution concentrations except for NO₂ at rural stations. Trends of the different pollutants (PM₁₀, PM_{2.5}, O₃, and NO₂) are discussed as well as exposure trends. The dataset covering the 2000–2015 period is available on a zenodo repository. In addition, the visualization of the maps is also available on the INERIS website with the possibility to download the data for more recent years. The presented dataset is of interest for air quality community, for example for comparison of air pollution trends in different countries, the dataset providing information for France. It is then suitable for publication, but some major issues should be addressed before publication (see point 1 and 2 of main comments):

Main comments:

The description of the kriging approaches is very limited in the paper and most of the references provided by the authors are written in French, limiting the access to non-French speaking readers. Providing a more detailed summary of the approaches would be valuable for the readers. The presented dataset is a fusion between model simulations and surface measurements.

We have given a more precise and detailed description of the kriging method used in this paper, as well as more extensive international references (see paragraph 2.3):

“Hourly atmospheric concentration fields are estimated by universal kriging, a geostatistical method. Kriging aims to estimate the value of a random variable (random process which describes the observations) at locations from the measurements. Kriging relies on the concept of spatial continuity which implies that measurements that are close to each other will be more similar than distant measurements. In addition, kriging requires a good knowledge of the spatial structure of the interpolation domain which is represented by the variogram or co-variogram (second order properties) of a random function (Goovaerts, 1997; Wackernagel, 2003; Chiles and Delfiner, 2012; Lichtenstern, 2013). Kriging involves deriving linear combination of the observations which ensures the minimal estimation variance under a non-bias condition. At a point s_0 , the concentration estimate $\widehat{y}(s_0)$ is given by equation 1.

$$\widehat{y}(s_0) = \sum_{i=1}^N \lambda_i y(s_i)$$

Equation 1

Where $y(s_i)$, $i=1\dots N$, are the observed concentrations at sampling locations through the entire domain (unique neighborhood) or within a limited neighborhood of s_0 (moving neighborhood), and λ_i , $i=1\dots N$, are the kriging weights.

Among the kriging methods, the universal kriging (especially external drift kriging) allows to consider additional information to make estimate more accurate. This approach is based on a linear regression

with auxiliary variables and a spatial correlation of the residuals and allows to combine simultaneously observations and additional information. The main hypothesis is that the global mean of the random variable is not constant through the domain and it relies on explanatory variables. This kriging technique has been used for several years in the monitoring air quality system for spatial interpolation at the regional scale (PREV'AIR, Malherbe et Ung, 2009). For $y(s_0)$, which is the pollutant concentration to be estimated at a location s_0 , the hypothesis is a linear relation between $y(s_0)$ and the considered auxiliary variables as explained by equation 2 and 3.

$$y(s_0) = m(s_0) + \varepsilon(s_0)$$

Equation 2

$$m(s_0) = b_0 + b_1x_1(s_0) + b_2x_2(s_0) + \dots + b_px_p(s_0)$$

Equation 3

Where $m(s_0)$ is the drift of the mean, b_0, b_1, \dots, b_p , are the coefficients of the linear regression, and x_0, x_1, \dots, x_p , are the auxiliary variables. ε corresponds to the stationary random process which is associated with a semi-variogram. In addition, the kriging weights must satisfy the drift condition described in equation 4.

$$\forall x_p : x_p(s_0) = \sum_{i=1}^N \lambda_i x_p(s_i)$$

Equation 4

In this work, kriging is performed with surface monitoring observations and the drift is described by the outputs from the CHIMERE chemistry transport model. European stations located outside the French domain are included in the kriging to increase accuracy at the borders. The kriging is performed using a moving neighbourhood as this allows for local adjustment of the relationship between the measurements and CHIMERE. The concentration at each grid point is estimated within a window of 80 monitoring sites. This number has been adjusted in previous studies by sensitivity tests (Benmerad et al., 2017; Beauchamp et al., 2017). In addition, smoothing is applied to avoid discontinuities in the map (Beauchamp et al., 2015b); the smoothing methodology was adapted from Rivoirard and Romary (2011). The final output resolution is the same as for the CHIMERE model: approximately 4 km resolution (0.06°×0.03°).

For PM10 (particles with a radius < 10 μm) and PM2.5 (particles with a radius < 2.5 μm) a co-kriging with external drift is applied. Co-kriging is an extension of kriging to the multivariate case. It allows the estimate of PM10 or PM2.5 concentrations by a linear combination of the two-variable data. The particularity of co-kriging is the use of the cross variance or semi-variance between the principal variable and the secondary variable. In the case of co-kriging with external drift, the simple and cross variograms are built based on residuals (Fouquet et al., 2007). Co-kriging allows to take into account the correlation between PM10 and PM2.5 and to improve consistency between PM10 and PM2.5 estimates (Beauchamp et al., 2015a). This cokriging also allows PM2.5 estimate to benefit from the higher density of PM10 monitoring stations.”

The authors do not provide any evaluation or discussion of the improvements provided by the kriging approaches compared to the raw model simulations. It would be very valuable to have this information to highlight the usefulness of the dataset compared to raw simulations. Is it possible to calculate the contribution of the model vs surface measurements for each grid point?

We calculated the validation scores for the raw data and added the following text to the paper on p. 16 (new section: 3.5: Comparisons with other scores):

“ In order to evaluate the added value of the kriging technique compared to the raw CHIMERE model simulations, the cross-validation scores can be compared to the raw model scores. Table 1 shows the scores averaged over all years and all background observations, without distinction of typology.

Table 1: Validation scores for the raw data and the kriged concentrations (cross-validation). Annual scores (bias, RMSE and the Pearson correlation coefficient r^2) are calculated over France for all year and all stations and are averaged.

	NO ₂	O ₃	PM ₁₀	PM _{2.5}
RAW				
Bias	-3.51	3.46	-8.91	-4.02
RMSE	12.97	17.26	12.63	8.73
R ²	0.55	0.73	0.71	0.75
KRIGED CONCENTRATION				
Bias	-0.51	-0.07	-0.04	-0.15
RMSE	10.41	12.54	7.64	5.83
R ²	0.81	0.92	0.85	0.87

All scores are clearly improved by the kriging of observations using CHIMERE as external drift. However, as can be seen in the previous figures, this improvement is more pronounced in urban areas than in rural areas, due to the much larger number of stations in urban areas, which makes the kriging more representative of these areas. “

The authors discussed the significance of the trends at the national scale, but few information is given when trend maps are presented. Are the trends significant at each grid point?

The representative confidence interval maps have not been included in the paper to avoid cluttering it up, but discussions of their results have been added.

Ex:

P22 (section 4.1): This trend is statistically significant on average over France with a narrow 95%-confidence interval ([-0.50;-1.09]) that does not include zero (see **Erreur ! Source du renvoi introuvable.**) and applies to almost all grid points (maps of confidence interval, not shown here)

p28 (4.1.3): When considering ozone, however, according to the value of the mapped 95 % confidence interval (not shown here) on most grid points, the confidence interval is wide and contains zero, indicating a lack of significance of the calculated trends.

A proofreading by a native English speaker is recommended.

Specific comments:

Page 2, lines 2-5: the authors should refer to the Tropospheric Ozone Assessment Report (TOAR activity from IGAC) when discussing tropospheric ozone trends.

The following reference has been added:

Tarasick, D., Galbally, I.E., Cooper, O.R., Schultz, M.G., Ancellet, G., Leblanc, T., Wallington, T.J., Ziemke, J., Liu, X., Steinbacher, M., Staehelin, J., Vigouroux, C., Hannigan, J.W., García, O., Foret, G., Zanis, P., Weatherhead, E., Petropavlovskikh, I., Worden, H., Osman, M., Liu, J., Chang, K.-L., Gaudel, A., Lin, M., Granados-Muñoz, M., Thompson, A.M., Oltmans, S.J., Cuesta, J., Dufour, G., Thouret, V., Hassler, B., Trickl, T. and Neu, J.L., Tropospheric Ozone Assessment Report: Tropospheric ozone from 1877 to 2016, observed levels, trends and uncertainties. Elem Sci Anth, 7(1), p.39. DOI : 10.1525/elementa.376, 2019

Page 4, CHIMERE description: the meteorological fields used as input of model simulations are different depending on the period (WRF from 2000 to 2010 and IFS from 2011). Does the change of systems to constrain the meteorological fields introduce any bias or discontinuity in the simulations?

It is indeed possible that the change in meteorological data between the period 2000-2010 and 2010-2015 has led to changes in the raw data. The evolution of the comparison scores of the raw model with the observation data seems to show higher correlations (r^2) after 2010 (not shown in the paper). However, it is difficult to know whether this can be attributed to meteorology alone since the emissions are also different. Furthermore, the WRF simulations themselves were nudged within ECMWF reanalyses, so they are not independent from IFS setup. Lastly, the data we produced are adjusted data using kriging methods. The impact of using either of the meteorological data sources will therefore be offset by the data fusion technique.

Figure 1 and similar: the dashed lines are confusing; they may be interpreted as error bars. They are not commented in the caption.

We added a description of those dashed lines in the figure captions.

Please check the size of the text in figures, it is sometimes too small, especially for figures from fig. 9.

Figures 9 to 15 have been enlarged.

Figure 9: the term "reanalysis" is used in the figure but never used in the text. Please use consistent terms all over the paper or define them clear (kriging, fusion, reanalysis).

The text has been made more consistent.